NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

SUITABILITY OF BOX-JENKINS MODELING
FOR NAVY REPAIR PARTS

by

Mark P. Businger

September 1996

Thesis Advisor: Robert R. Read

Approved for public release; distribution is unlimited.

19961220 111

DTIC QUALITY INSPECTED 1

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE September 1996	3. REPORT TYPE AND DATES COVERED Master's Thesis

4. TITLE AND SUBTITLE SUITABILITY OF BOX-JENKINS MODELING FOR NAVY REPAIR PARTS	5. FUNDING NUMBERS
6. AUTHOR(S) Businger, Mark P.	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey CA 93943-5000	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.	12b. DISTRIBUTION CODE

13. ABSTRACT *(maximum 200 words)*

A basic function in the proper management of repair part inventories is the forecasting of future demand. The Navy maintains a database of univariate demand data for its repair part inventories using a quarterly time interval. Historically, Navy repair part demand forecasting has been done using the exponential smoothing procedure. This method is a simple and robust means of forecasting, however it does not make use of any characteristics of the entire time series such as trend, cycles, presence of outliers, or demand clustering. This research begins by developing several simple, robust, and dimensionless time series features. These features are used to predict the suitability of Box-Jenkins (ARIMA) modeling. The ARIMA process is a powerful time series modeling and forecasting technique which possesses flexibility for the inclusion of many time series characteristics. This research project develops a predictive model of ARIMA suitability using both classical regression and a modern expert-system statistical package, ModelQuest. A computationally simple means is presented for determining which time series may benefit from the Box-Jenkins methodology. Using ARIMA modeling for time series that show significant benefit will provide a more accurate demand forecast and benefit inventory management.

14. SUBJECT TERMS Time Series, Box-Jenkins, Forecasting, Univariate, Inventory	15. NUMBER OF PAGES 72
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18 298-102

Approved for public release; distribution is unlimited.

SUITABILITY OF BOX-JENKINS MODELING FOR NAVY REPAIR PARTS

Mark P. Businger
Lieutenant, United States Navy
B.S., Stanford University, 1986

Submitted in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 1996

Author: Mark P. Businger
Mark P. Businger

Approved by: R R Read
Robert R. Read, Thesis Advisor

Lyn R. Whitaker
Lyn R. Whitaker, Second Reader

Frank C. Petho
Frank C. Petho, Chairman
Department of Operations Research

ABSTRACT

A basic function in the proper management of repair part inventories is the forecasting of future demand. The Navy maintains a database of univariate demand data for its repair part inventories using a quarterly time interval. Historically, Navy repair part demand forecasting has been done using the exponential smoothing procedure. This method is a simple and robust means of forecasting, however it does not make use of any characteristics of the entire time series such as trend, cycles, presence of outliers, or demand clustering.

This research begins by developing several simple, robust, and dimensionless time series features. These features are used to predict the suitability of Box-Jenkins (ARIMA) modeling. The ARIMA process is a powerful time series modeling and forecasting technique which possesses flexibility for the inclusion of many time series characteristics. This research project develops a predictive model of ARIMA suitability using both classical regression and a modern expert-system statistical package, ModelQuest. A computationally simple means is presented for determining which time series may benefit from the Box-Jenkins methodology. Using ARIMA modeling for time series that show significant benefit will provide a more accurate demand forecast and benefit inventory management.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This thesis explores a means of determining the suitability of Box-Jenkins, or Auto-Regressive Integrated Moving Average (ARIMA), modeling for the demand forecasting of Navy repair part data. The ARIMA process is a powerful method of data fitting and forecasting that is sensitive to patterns or processes within univariate time series. This process is a desirable alternative to other less powerful but computationally simpler methods such as exponential smoothing, which is currently in use by the Navy. The primary drawback to the use of ARIMA modeling is the increased complexity of the model, the intensity of the computation, and the fact that many parts do not benefit from its use. It is estimated that 32% of Navy repair parts with recurring demand will show significant benefit from ARIMA modeling.

A method of evaluating which time series will benefit most by ARIMA modeling is developed. Several computationally simple features are presented and applied to the time series. These features seek to identify the presence of different and relevant characteristics within the individual time series. Examples of these features include: trend, skewness, seasonality, and autocorrelation. A method of classifying the specific time series is generated that takes these features compositely. The suitability of ARIMA to a particular demand stream is assessed by processing the set of features through a classical regression and an expert-system statistical network program. These methods create a fast and simple means of discriminating which time series will benefit most, and are therefore worthy of, ARIMA modeling.

ACKNOWLEDGEMENT

The author wants to thank Dr. B. Bissinger and Mr. J. Boyarski for their guidance in laying the foundation for this research. The work performed in this thesis benefited greatly from their insight into time series processes and characteristics.

I. INTRODUCTION

The Navy Supply System's primary responsibility is stocking, maintaining, and providing the spare parts essential to support naval ships and shore establishments. Proper management of these spares is necessary to maintain operational effectiveness. Making the right choices with regard to which repair parts to stock and at what levels to stock them gains importance as the number of ships and their support base decreases.

A. REPAIR PART FORECASTING

The Naval Inventory Control Point at Mechanicsburg, PA (NAVICP-Mech) is responsible for managing the repair part inventories required by the Navy. Demand forecasting is the basic building block of this inventory management. Using historical demand data, forecasts are developed for each repair part. These forecasts are then used to derive inventory stocking levels. Historically, the method used to generated these forecasts has been exponential smoothing. This technique uses a simple weighted average between the forecast for the current period and the actual demand for the current period to come up with a forecast for the next period. The primary advantages of the exponential smoothing process are its simplicity and the low computer data storage requirement.

Over the past 10 years, however, the aggregate quarterly demand data has been kept in a separate database. For each repair part, a single number representing the total quarterly demand from all activities was recorded and stored. The aggregate of quarterly

demand observations for each repair part forms a univariate time series. Typical examples of these time series are shown in Figure 1.

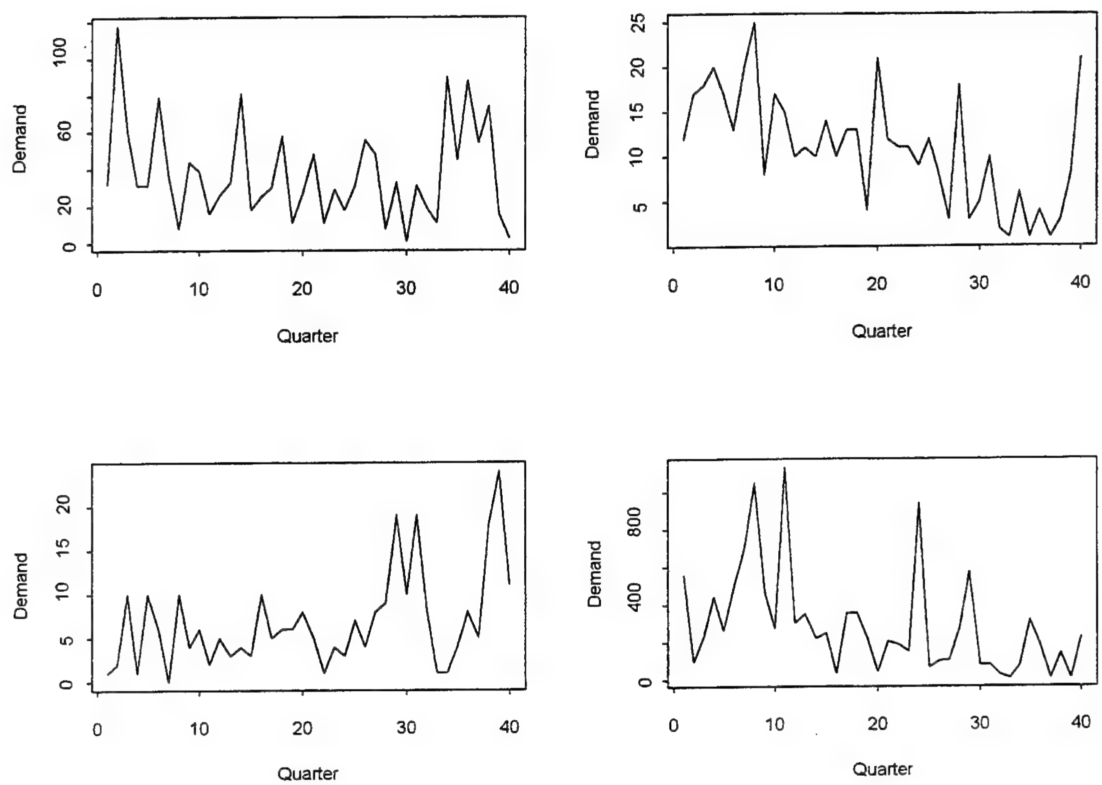


Figure 1: Examples of Repair Part Demand Histories

The primary purpose of this database is to capture a demand history, or demand stream, for possible future analysis. No covariates of demand have been developed nor recorded, however these are certainly worthy of consideration for selected instances in future analyses. Possible covariates include: number of ships or number of shore

part is employed, the criticality of the repair part, the size or handling difficulty of the part, the shelf life, and the mean age of the supported equipment.

The goal of this thesis is to examine and explore the characteristics, or features, of these univariate demand streams and develop measures of suitability for Box-Jenkins modeling. A means of categorizing the demand process emerges upon studying the features of a demand stream. This thesis explores various features with respect to their ability to both quantitatively and qualitatively anticipate the suitability of Box-Jenkins, or auto-regressive integrated moving average (ARIMA), modeling.

B. THESIS ORGANIZATION

Chapter II provides background regarding exponential smoothing, the time series features, and the Box-Jenkins methodology. Chapter III contains the motivating methodology of the thesis. It describes the foundations for the research, and presents details regarding the calculations performed prior to the analysis. Chapter IV contains the analyses, both quantitative and qualitative, and presents classification rules for determining ARIMA modeling suitability. Finally, Chapter V provides recommendations and conclusions.

II. BACKGROUND

NAVICP-Mech is currently responsible for a line item inventory of over 230,000 stock numbered repair parts. A large number of these repair parts are inactive and have experienced no recent demand, though. For this thesis, a database of the demand histories for 139,491 repair parts, comprising those with recent active demand, was extracted for study.

A. EXPONENTIAL SMOOTHING

Exponential smoothing is a technique that incorporates knowledge of the entire history of a process to form a forecast of the next observation. It is commonly used as a simple and rational means of time series forecasting. The equation for exponential smoothing is

$$S_t = \alpha\, y_t + (1-\alpha)S_{t-1} \qquad 0 < \alpha < 1 \tag{2.1}$$

where S_t is the forecast made in period t, α is the exponential smoothing coefficient, and y_t is the demand in period t. Recursively substituting into equation (1) results in

$$S_t = \alpha\, y_t + (1-\alpha)\alpha\, y_{t-1} + (1-\alpha)^2\alpha\, y_{t-2} + \ldots + (1-\alpha)^{t-1}\alpha\, y_1 \tag{2.2}$$

This result shows that the forecast for any time t is a linear combination of all previous demands observed, with increasing weight given to more recent observations.

The initial impetus for using exponential smoothing to forecast repair part demand was the low requirement of computer data storage and simplicity of calculations. Strict boundaries on amount of data archivable in the early computers ruled out other choices.

B. DEMAND STREAM CHARACTERISTICS

The shortfall of exponential smoothing is that it does not consider characteristics of the demand stream as a whole. There are possibly both long and short term processes at work in a time series. For example, features such as trend, cycle, or discrete shifts in the process' mean or variation may be characteristic of a particular demand stream. The characteristics of each demand stream can uniquely describe and categorize it. Bissinger and Boyarski (1992) present a detailed discussion of the possible patterns in that may occur. This thesis develops a basic ability to evaluate the presence of certain distinct features. The suitability of alternative modeling methods such as the Box-Jenkins methodology can be determined using this ability.

C. BOX-JENKINS FORECASTING

Two powerful time-series modeling processes combined by Box and Jenkins are moving average (MA) and autoregression (AR) (Box and Jenkins, 1970). These processes attempt to separate the time series process into its components of an underlying process and a white noise process. The white noise process contributes the unpredictable random element to the time-series process, while the underlying process can be examined and used for analysis and prediction.

A moving average process of order q is defined by

$$X_t = -(B_1 E_{t-1} + \dots + B_q E_{t-q}) + E_t \tag{2.3}$$

where E_t is an unobservable error term and the $\{B_i\}$ represent the moving average parameters. The interpretation of the MA process of order q is that any given value in the series, X_t, is directly proportional to the q previous random errors, $\{E_i\}$, plus some current random error E_t. This means that the prediction for period t is based only on the random errors that have occurred in the q previous periods.

An autoregressive process of order p is one which recursively satisfies

$$X_t = A_1 X_{t-1} + A_2 X_{t-2} + \dots + A_p X_{t-p} + E_t \tag{2.4}$$

where E_t is again an error term representing a process with mean of zero and a finite variance and the $\{A_i\}$ are the auto-regressive parameters. An autoregressive model of order p expresses a time series value as the arithmetic combination of p past series values plus a random error term. With $p=1$, the autoregressive process is identical to an exponential smoothing process (Box and Jenkins, 1970).

Combining these two processes into one with both MA and AR parameters results in a model of a stationary autoregressive moving average process (ARMA). A stationary process has a constant underlying mean. The ARMA model represents any series value as the combination of both past series values and past random error values. The assumption of stationarity is frequently an inappropriate one, though, and the regular ARMA process may therefore be inadequate. Stationarity can often be induced by using a differenced demand history. A single differenced series is attained through the equation:

$$Z_t = X_t - X_{t-1} \tag{2.5}$$

This differencing operation is integrated into the ARMA process by substituting the differenced values, Z_i, for the original values. The difference process may be repeated as often as necessary to provide stationarity. The order of the differencing operator is denoted by d.

The inclusion of the differencing operator to the ARMA process results in an autoregressive integrated moving average (ARIMA) model. The ARIMA(p,d,q) model combines several powerful approaches to time-series analysis into a single package. The process is defined by all three parameters: the order of the autoregression, p, the order of the integration operator, d, and the order of the moving average process, q.

The difficulty in applying the ARIMA process to the NAVICP-Mech demand streams is two-fold. First, the time series available for analysis are short ones. The ARIMA process works best over long times series that exhibit stationarity, but in the repair part database there are a maximum of 40 time intervals available for analysis. However, maintaining a more lengthy history will not likely assist the modeling process in this case. The database already contains 10 years worth of data – a long period of time when considering the total life cycle for a supported system. The addition of demand measurements that are over 10 years old will likely provide little insight into the present demand process.

The second difficulty is the necessary computer time. Since ARIMA will not provide significant benefits in every instance, it is prohibitive to calculate directly the

suitability for all repair parts. Using the statistical software package S-Plus® on a typical personal computer system running at 66 Mhz, the calculation of a proper ARIMA model to fit a single demand stream requires slightly over 10 seconds of processing time. Using a large mainframe computer will certainly reduce the processing time greatly, however the time required to process a multitude of demand streams will still be quite large. A better way to identify stock numbers which benefit by ARIMA modeling must be found in order to take advantage of the modeling benefits it affords.

III. METHODOLOGY

Since it is infeasible to conduct an analysis of the entire data set provided by NAVICP-Mech, we first develop a logical methodology of sampling and evaluation. The initial task in defining the methodology is the specification of a sampling plan. The one used in this thesis begins by partitioning the data into subsets in which some degree of homogeneity can be expected. Two measures are selected for partitioning: the coefficient of variation and the number of zeros within the demand stream. Once these cross classified subdivisions are established, each one is sampled at random to provide demand streams for detailed analysis.

A measure of effectiveness (MOE) to describe ARIMA modeling suitability is then defined and applied to all the sampled demand streams. A means of predicting that same effectiveness is developed next. This prediction is based upon the characteristic features of each of the demand streams. Each feature is defined based on its ability to measure and describe a basic characteristic.

A. SAMPLING SCHEME

Not surprisingly, a large portion of the demand streams and the database as a whole consists of zeros. This is due to the high percentage of repair part items which are maintained as insurance items but which rarely experience demand. Demand streams with a high number of zeros are exceedingly difficult to analyze. For the purpose of this thesis, a high total number of zeros is viewed as a process with an intermittent random

component. Only those demand streams with at least 20 positive demands were considered for analysis. After the demand streams with high zero counts are filtered out, the set of interest, N_o, comprises 12,000 demand histories.

The 12,000 observations in N_o are viewed in several different ways in search of clear partitions. These partitions are chosen to define a sampling scheme, and possibly to provide an initial insight into the distribution of stock numbers by feature. The work of Boyarski (1995) suggests several different measures to be used in this initial analysis. Specifically, the following characteristics are identified as very meaningful: demand mean, demand stream length, coefficient of variation, and number of zeros within the demand stream. Additional measures that were considered were mean length of non-zero runs, mean length of zero runs, median, and the ratio between the mean and the median.

The distribution in the data for each of these measures was examined and graphed, both singly and in bivariate groups, in search of defensible partition values. While most characteristics had distributions that displayed a sharp right skewness and no clear divisions, the coefficient of variation distribution was unimodal, surprisingly symmetric, and centered rather close to one.

Based on the encouraging distribution of the coefficient of variation, it was chosen as the primary criteria for developing a means to partition N_0. The quantile breakpoints for the 12,000 element subset were evaluated and are shown in Table 1.

Quantile	Coefficient of variation value
0.2	0.8
0.4	1.0
0.6	1.2
0.8	1.5

Table 1: Empirical quantiles of dataset N_0

The five groups resulting from separation by coefficient of variation (c.v.) quantile are then examined further. A distinctive feature which provides for a large degree of variation within the quantile subsets is the number of zeros within the demand stream, and this feature is chosen as a secondary discriminator.

In using this discriminator, it is again desirable to have roughly equal numbers of stock numbers in each bin. Since the spread of the number of zeros varies greatly between the different c.v. quantile subsets, secondary breakpoints are chosen in each of the subsets. Table 2 shows the final partitioning, divided by range of number of zeros within the time series.

Subclass:	Class: A $0 \le \frac{\sigma}{\mu} < 0.8$	B $0.8 \le \frac{\sigma}{\mu} < 1.0$	C $1.0 \le \frac{\sigma}{\mu} < 1.2$	D $1.2 \le \frac{\sigma}{\mu} < 1.5$	E $1.5 \le \frac{\sigma}{\mu}$
1	0	0-2	0-5	0-8	0-8
2	-	3-5	6-9	9-12	9-12
3	1	6-7	10-12	13-16	13-16
4	2-3	8-10	13-15	17-18	17-18
5	4-20	11-20	16-20	19-20	19-20

Table 2: Divisions of N_o by Number of Zeros within the Demand Stream

An identification scheme for each of the divisions is developed. Each cell is identified by two characters. The class, an alphabetic character A through E, corresponds to the divisions by coefficient of variation. The subclass, a numeric character 1 through 5, is assigned based on grouping by number of zeros within the class.

A sample of 45 demand streams was taken from each subset of N_0. This represents roughly a 10% sample from each division, with a total of 1080 demand streams sampled. The analysis portion of this thesis relies on these subsets to create and test statistical models.

B. MEASURE OF EFFECTIVENESS (MOE)

The base MOE adopted for a time-series model is the ratio of the standard deviation of the demand stream to the mean squared residuals for the predictive model. Algebraically,

$$MOE = \frac{s}{\sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n-1}}} \tag{3.1}$$

An effective modeling process is one for which this ratio is above one.

C. FEATURES

Several demand stream features are developed and implemented in the course of this thesis. Of primary importance while developing these features are the considerations of computational complexity, robustness, and dimensionality. A very low computational complexity is essential in making the calculation of the features an inexpensive process. Robustness implies that the features can both be calculated almost universally and will provide meaningful values over all of the different demand streams. Dimensionless measures are desirable since their resultant values can be easily compared from demand stream to demand stream. When taken compositely, these measures provide for the classification of a particular repair part.

1. Coefficient of Variation

The coefficient of variation is defined as $s/\bar{x}$. By measuring the standard deviation as a fraction of the mean, the coefficient of variation evaluates how "noisy" a process is. Larger values of the ratio correspond to increasingly more random, or noisy, processes.

2. Number of Zeros

The number of zeros is simply the count of zeros within the time series. In the NAVICP-Mech database, there are a large number of demand streams that begin with one (or more) zeros. These leading zeros were not considered when counting the number of zeros. For the purpose of this thesis, the count of number of zeros begins at the first non-zero entry in the time series.

3. Trend

Trend is the component of a demand stream that measures the rate of growth or decline over time. The feature of trend seeks to quantify the existence and degree of the trend component within a series. The development of this feature begins with considering the demand series in the form (t_i, y_i) where $i = 1,\ldots n$, t_i denotes the specific quarter and y_i is the demand for the quarter. The demand series is then divided into thirds according to time. This means that $n = 3k+r$ where k is a positive integer and r is the remainder term, $r = 0,1,2$. Table 3 displays the different divisions possible.

Remainder term	First third	Middle third	Last third
$r = 0$	k	k	k
$r = 1$	k	$k + 1$	k
$r = 2$	$k + 1$	k	$k + 1$

Table 3: Series divisions based on remainder

For the pairs in the lower third, (t_i, y_i), where $i = 1,\dots k$, let m_{t1} and y_{t1} denote the medians. Similarly, for the pairs in the upper third, (t_i, y_i), where $i = n\text{-}k\text{+}1,\dots n$, let m_{t3} and y_{t3} be the respective medians as well.

From these results, a robust line (McNeil, 1977) is calculated by

$$\frac{y_{t3} - y_{t1}}{m_{t3} - m_{t1}} \tag{3.2}$$

In order to convert this to a dimensionless index, the demands are first ordered such that

$$y_{(1)} \leq y_{(2)} \leq \dots\dots \leq y_{(n)} \tag{3.3}$$

Then, the *100/6* and *500/6* quantiles are estimated. These values are denoted by $y_{(1/6)}$ and $y_{(5/6)}$ respectively, and play the role of m_{t1} and m_{t3} respectively. The dimensionless trend index then is given by

$$Trend\ Index = \frac{y_{t3} - y_{t1}}{y_{(5/6)} - y_{(1/6)}} \tag{3.4}$$

For the trend index, values close to zero support no trend, while values close to +/- 1 support trend of the appropriate sign.

4. Peaks

Peaks are outliers of the demand process. A peak represents a significant and abrupt deviation from the established process. A robust measure of peaks was developed for this thesis. It utilizes a 10% trimmed mean and variance. Using the trimmed mean and variance provides for a magnitude resistant method of measuring outliers. These measures are given by

$$\bar{y}_{t\,(0.1)} = \frac{1}{n_{0.8}} \sum_{n_{(0.1)}}^{n_{(0.9)}} y_{(i)} \tag{3.5}$$

$$s^2_{t\,(0.1)} = \frac{1}{n_{0.8} - 1} \sum_{n_{(0.1)}}^{n_{(0.9)}} (y_{(i)} - \bar{y}_{t\,(0.1)})^2 \tag{3.6}$$

where $n_{0.8}$ represents the number of observations used, and $n_{(0.1)}, n_{(0.9)}$ represent the first and last subscripts used within the trimmed series

The measure of peaks is measured in terms of the number of standard deviations from the mean. The measure of peaks is made relative to the standardized value

$$r_i = \frac{\left| y_i - \bar{y}_{t\,(0.1)} \right|}{s_{t\,(0.1)}} \tag{3.7}$$

By Normal theory, it is expected that about 5% of these measures within a series will be larger than 2 and these members are defined as peaks . The feature of interest is the total count of the number of peaks within a series.

5. Seasonality

Seasonal variations consist of regular cycles which occur when demand fluctuates in a repetitive pattern. The most evident cycle in every day usage is an annual one, and the feature of seasonality in this thesis gauges the degree to which repair part demand fluctuates according to the four seasons within one year. Other periods of seasonality may prove useful, and can be calculated easily through simply modifying the method described below.

A winsorized time series is used in order for this feature to be resistant to outliers. The winsorization process replaces the tail order statistics so that outliers do not provide excessive influence (Miller, 1986). At the same time, it prevents the data loss associated with trimming. The winsorized time series is given by

$$X_i = \begin{cases} X_{n_{(0.1)}-1}, & i = 1,\ldots,n_{(0.1)}-1 \\ X_i, & i = n_{(0.1)},\ldots,n_{(0.9)} \\ X_{n_{(0.9)}+1}, & i = n_{(0.9)}+1,\ldots,n \end{cases} \tag{3.8}$$

Our feature of seasonality requires that the series be distributed among four bins. Let $n = 4k + r,\ \ r = 0,1,2,3$. Then the winsorized series is separated into four bins, y_{1j} ,y_{2j} ,y_{3j} and y_{4j}, with each containing k or $k+1$ entries. A series element is a member of bin $y_{i,j}$ when it can be written as $y_{i,i+4(j-1)}$ for $i = 1,2,3,4$ and $j = 1,2,\ldots k$. For each of the bins, $\bar{y}_i$ and $\bar{s}_i^2$ is calculated. In effect, a one-way analysis of variance (ANOVA) is performed to produce the within and total sums of squares. The coefficient of determination is given by

$$R^2 = 1 - \frac{\sum_{i=1}^{4} SumOfSquares_{within\ bin}}{SumOfSquares_{total}} \tag{3.9}$$

This serves as the measure of seasonality, where values close to one indicate high seasonality, while values close to zero indicate no seasonality.

6. Runs

The feature of runs measures the tendency within a demand stream to have consecutive observations either above or below the median of the overall process. A run is defined as a collection of successive observations all on the same side of the process median. This measure of runs is useful in representing the amount of clustering of observations that occurs.

The clustering property is evidenced by groupings, or clumps, of observations interspersed by periods of no or little demand (Bissinger and Boyarski, 1992). Small values of the runs feature correspond to a tendency to cluster while large values indicate a more oscillatory process. This feature should always be considered in conjunction with the trend feature, since a large trend component may lead to a false indication of no clustering.

In measuring runs, first calculate the median of the demand stream. If any demand observations are equal to the median, they are discarded for the computation. The number of runs above and below the median are then counted. These are denoted by R_1 and R_2, respectively. The sum of these two measures is denoted by R. Two additional necessary

measures are n_1, the total number of observations above the median, and n_2, the total number of observations below the median.

According to Gibbons (1992), the expected value and variance of R are given as

$$E[R] = 1 + \frac{2\,n_1\,n_2}{n_1 + n_2}, \tag{3.10}$$

$$Var(R) = \frac{2\,n_1\,n_2(2\,n_1\,n_2 - n_1 - n_2)}{(n_1 + n_2)^2\,(n_1 + n_2 - 1)} \tag{3.11}$$

The feature of runs can then be standardized into a dimensionless measure

$$Runs = \frac{|R - E[R]|}{\sqrt{Var(R)}} \tag{3.12}$$

7. Skewness

Skewness is the tendency of a distribution to be asymmetric. To evaluate skewness, the comparison is made between the mean and the median of a series. Simply, an index of skewness is given by

$$I_{sk} = \bar{y} / m_y, \tag{3.13}$$

where $\bar{y}$ is the arithmetic average demand and m_y is the median demand. For I_{sk} values greater than one, the process is skewed to right. Conversely, values less than one indicate a process skewed to left.

8. Autocorrelation

Autocorrelation is a measure of how strongly time series values are correlated to each other over time. A nonparametric measure of the overall autocorrelation based on a single lag element is chosen. The measure presented by Gibbons (1992) is used.

First, assign each observation y_i a rank r_i relative to all the other elements of the series. That is to say, the lowest y_i is assigned a rank of one and the largest is assigned a rank of n. Then form the following measures:

$$NM = \sum_{i=1}^{n-1} [r_i - r_{i+1}]^2, \tag{3.14}$$

$$RVN = \frac{NM}{\sum_{i=1}^{n} [r_i - \frac{(n+1)}{2}]^2} \tag{3.15}$$

If there are no ties within the ranks, this simplifies as

$$RVN = \frac{12\ NM}{n(n^2 - 1)} \tag{3.16}$$

The expected value of this measure is 2, and its variance is given by

$$\sigma^2 = \frac{4(n-2)(5n^2 - 2n - 9)}{5n(n+1)(n-1)^2} \tag{3.17}$$

This allows RVN to be normalized according to

$$RVN_{norm} = \frac{|RVN - 2|}{\sigma} \tag{3.18}$$

This result is the feature of autocorrelation, with small values indicative of positive autocorrelation and large values indicating negative autocorrelation.

IV. ANALYSIS

The characteristic features are first evaluated and stored for each of the 1080 elements of the dataset. The program to calculate the features and summarize them into a convenient database was written in Turbo-Pascal. When compiled, the Turbo-Pascal code is extremely efficient in terms of speed of calculation. Less than 30 seconds of total computational time are required to calculate and store the values for all of the features for the entire dataset.

The MOE of ARIMA suitability is calculated using the S-Plus® statistical package. The MOEs for exponential smoothing as well as for three different orders of ARIMA models were calculated. Due to the large number of different ARIMA models available, the three models were chosen to provide as great a range as possible. The three models chosen are: a first order (1,1,1) model, a second order (2,2,2) model, and a third order (3,2,3) model. In the third order model, the order of integration is left as two since integrations of order three (or higher) are almost never necessary (Hoff, 1983). The results are shown in Figure 2.

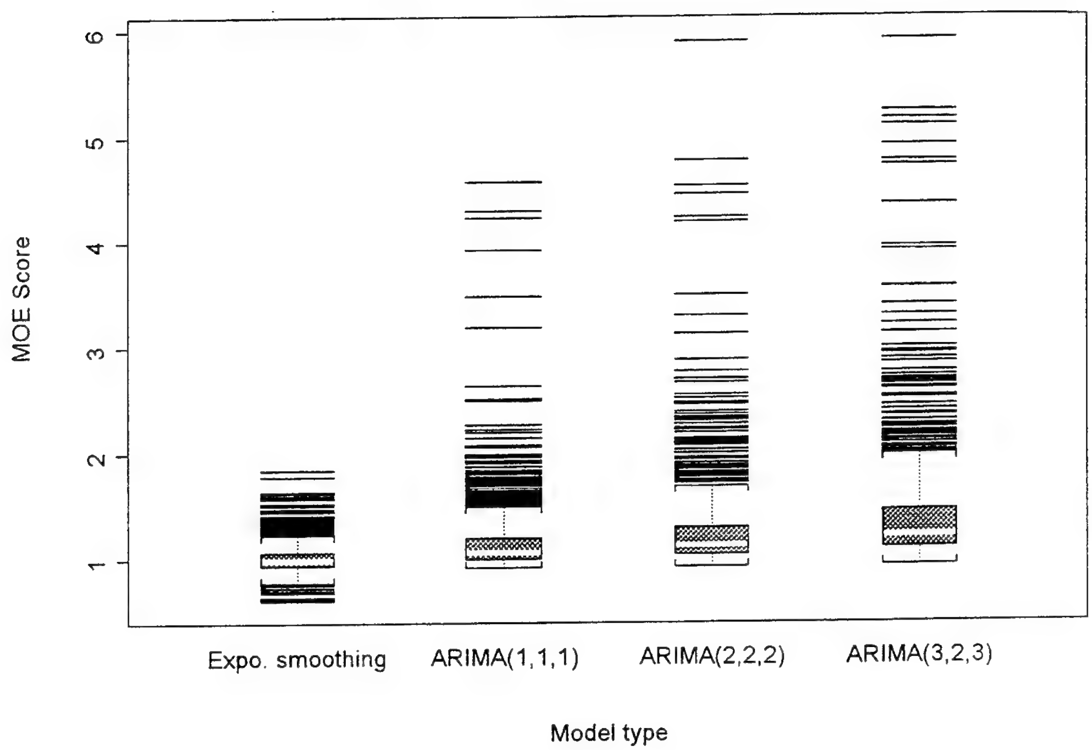


Figure 2: Boxplots of MOE Values for Four Different Modeling Techniques

Figure 2 shows that there are a number of demand histories for which ARIMA modeling provides an appreciable benefit over exponential smoothing. At the same time, however, there are a number of demand histories for which exponential smoothing provides nearly as much predictive power as any of the other models. The mean MOE improvement when compared to exponential smoothing is: 0.15 for the ARIMA(1,1,1) model, 0.28 for the ARIMA(2,2,2) model, and 0.41 for the ARIMA(3,2,3) model.

While it would be desirable to have MOE improvement in separate types of time series using separate levels of ARIMA modeling, such does not appear to be possible. Comparison of the MOE results shows that there was one general group of time series which benefited from ARIMA modeling. Where ARIMA(1,1,1) provides an appreciable benefit over exponential smoothing, higher order ARIMA models generally provide an additional benefit over the single order model. In those cases where ARIMA(1,1,1) provides little benefit, the other models usually provide little improvement as well. This observation is evidenced again in the statistical network analysis.

A. STAR PLOTS

A better understanding as to which features of a demand history are indicative of ARIMA model suitability is desired. An initial means of exploring this is through star plots as described by Chambers, Cleveland, Kleiner, and Tukey (1983). A star plot converts multi-attribute data into graphical stars. The lengths of each specific ray of the star corresponds to the magnitude of that attribute relative to the same attribute within the other stars. A pattern may emerge of features that are indicative of model suitability by grouping star plots by MOE level and comparing those demand histories that benefited from ARIMA modeling. A diagram of the star plot format and examples of star plots with likely ARIMA modeling benefit are shown in Appendix A.

Examination of the star plots shows that several of the features appear to have ARIMA predictive power. A large measure of autocorrelation paired with an indication of trend is a common element in many of the demand histories that benefited most from

ARIMA. Coefficient of variation does not seem to play a significant role. A moderate measure of skewness was usually evident in the histories with larger MOEs. Several of the star plots were difficult to visually identify as ARIMA candidates, due to the lack of a definitive common element. There is also a wide disparity between the star plots of those histories which experience a moderate improvement in the MOE and those with little improvement.

The inclusion of additional features in the star plots was explored on the computer. The features of variance, mean, mean length of zero runs, and total number of observations within the demand stream were added. None of these additional features provide any improvement to the visual categorization of the star plots, and often detracted from the categorization. The additional features were subsequently removed from further consideration.

An interesting result from the examination of the star plots is that the samples taken from the E class, representing those with a coefficient of variation greater than 1.5, have little similarity with those from the smaller coefficient of variation subsets. Elements from the E class with a high MOE measure have a variety of star plot shapes, and are generally very different from those with a high MOE from the other subsets. This suggests that there may be an undiscovered process at work in the high coefficient of variation demand streams that is obscuring ARIMA model suitability.

B. CLASSICAL REGRESSION

The standard method for determining the relationship between predictor variables, namely the demand features, and resultant variable, the MOE, is through classical regression. Regression analysis is conducted using the eight explanatory variables from the star plots: coefficient of variation, number of zeros, trend, peaks, seasonality, runs, skewness, and autocorrelation. The first regression tested is a simple additive model of the explanatory variables, as shown in the Equation 4.1. Initially, the MOE values for ARIMA(1,1,1) models are used as a resultant measure.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7 \qquad (4.1)$$

where y is the MOE measure,
x_1 is the coefficient of variation,
x_2 is the number of zeros within the series,
x_3 is the trend feature,
x_4 is the measure of peaks,
x_5 is the seasonality feature,
x_6 is the measure of runs,
x_7 is the skewness feature,
x_8 is the autocorrelation measure,
and β_i are the regression coefficients

The regression results for the data set predicting ARIMA(1,1,1) suitability are very modest. The resultant p-value of 0.0 indicates strong evidence of a relationship between the predictors and the resultant, but the R^2 value of 0.16 implies that only a small fraction of the MOE can be successfully predicted.

The post-regression analysis provides a likely reason for this poor performance. Examination of the raw residuals shows several that are outlier candidates. After

calculating their Studentized residuals and Cook's distances, it is clear that these are highly influential outlier points. Interestingly, the outliers are primarily grouped together within the E class. Further investigation indicates that almost all the elements of the E class had significantly larger Cook's distances compared to the A through D classes. Based on this result, and the results from the qualitative analysis of the star plots which show large variability within the E class, a structurally identical regression is performed using classes A through D only.

The results of the second additive regression are very promising. The p-value remains at 0.0, while the R^2 value increases to 0.51. The post-regression analysis reveals a lingering bow in the residuals vs. fitted values plot, but otherwise no surprising results. Excluding the class with the largest coefficient of variation results in a significant improvement in predictive power. It is evident that the E class will require special treatment, since it behaves noticeably differently than the other subsets. For the purposes of this thesis, modeling will be done using only the classes A through D.

Consideration was given as to whether the response variable, MOE measure, should be transformed to provide a better fit. The primary indication that a better fit might be obtained is the bow in the residuals plot. We make use of the quantitative method Box and Cox (1964) developed to find the optimal transformation. After fitting a full model to the data, the Box-Cox transformation algorithm is applied to the model to obtain a 95% confidence interval for the optimal transformation. Based on this result, a power transformation with $\lambda = -3.5$ is applied to the full model. This results in a p-value of 0.0

and an R^2 value of 0.65. The R^2 values between the transformed and untransformed models cannot be directly compared, since they model different responses. However, the results of the transformed regression can be "untransformed" and a comparative R^2 generated using the ratio of the within sum of squares to total sum of squares. This calculation results in an R^2 of 0.68, representing an average variance reduction from the original R^2 of 0.51. The post-regression analysis shows a much improved distribution of residuals with the transformed model as well.

A determination as to which, if any, of the predictor variables are extraneous in the model is appropriate in validating the model. A stepwise regression process is performed. This thesis uses Efroymson's stepwise method, which is similar to forward selection. The significant difference with forward selection is that after each predictor addition, partial correlations are considered to see if any of the variables in the subset should now be dropped. By either adding or subtracting terms at each step, an efficient model with a minimal number of terms results. All the single variable terms as well as the second order interaction terms between predictors are included in the stepwise regression. The interaction terms are significant in that they consider the relationships between predictive variables. From the stepwise regression, the final regression model for predicting single order ARIMA coefficients is

$$\begin{aligned} y^{-3.5} = {} & \beta_0 + \beta_1 x_3 + \beta_2 x_5 + \beta_3 x_6 + \beta_4 x_1 x_3 + \beta_5 x_2 x_3 + \beta_6 x_2 x_5 + \\ & \beta_7 x_2 x_8 + \beta_8 x_3 x_8 + \beta_9 x_4 x_7 + \beta_{10} x_5 x_6 \end{aligned} \tag{4.2}$$

where y is the MOE measure,

x_1 is the coefficient of variation,

x_2 is the number of zeros within the series,

x_3 is the trend feature,
x_4 is the measure of peaks,
x_5 is the seasonality feature,
x_6 is the measure of runs,
x_7 is the skewness feature,
x_8 is the autocorrelation measure,
and β_i are the regression coefficients.

Noteworthy in these results is the fact that all of the features previously developed contributed to the overall regression equation, either singly or as a part of an interaction term. The untransformed R^2 value for the stepwise regression result is 0.7734. The detailed results of the regression for the ARIMA(1,1,1) resultant are shown in Appendix B.

Graphing the tranformation inverted fitted values vs. their residuals shows an interesting result. Resultant values at the high-end of the MOE scale, above approximately 1.6, are predicted to be significantly higher than they actually are. This result is due to the large power of the transformation, which skews the values when the transform in inverted. With this knowledge, actual predictions from the regression model would have an effective R^2 greater than the evaluated untransformed one of 0.7734 since a compensation can be made to larger values of predicted MOE. The plots of transformation inverted fitted values vs. residuals are shown in Appendix C for all three cases that were ARIMA modeled.

A similar regression analysis as with the single order ARIMA is conducted using the ARIMA(2,2,2) and ARIMA(3,2,3) MOE values and calculated features. The method of analysis is the same as for the ARIMA(1,1,1) case. In brief, the simple additive model

with an untransformed result is first developed. A transformation for the result variable, using the Box-Cox algorithm, is then considered. Finally, a stepwise regression is conducted using the transformed result variable and all the predictor variables as well as their interaction terms as explanatory variables. At all steps in the regression process, the results are critically evaluated for influential points, outliers, or other regression anomalies.

For the ARIMA(2,2,2) case, the final equation is of the form

$$\begin{aligned} y^{-3.0} = {} & \beta_0 + \beta_1 x_3 + \beta_2 x_5 + \beta_3 x_6 + \beta_4 x_8 + \beta_5 x_2 x_3 + \beta_6 x_3 x_8 + \\ & \beta_7 x_4 x_6 + \beta_8 x_5 x_6 + \beta_9 x_5 x_8 \end{aligned} \tag{4.3}$$

The regression equation for the ARIMA(2,2,2) case results in an R^2 of 0.6908 when the transformation is inverted. Inspection of the plot of untransformed fitted values vs. untransformed residuals shows a similar result as with the ARIMA(1,1,1) case. There are three points at the high-end of the MOE values which contribute the bulk of the error in prediction. As a result of the powerful transformation in Equation 4.3, these values are predicted to be significantly higher than their actual values. The plot of untransformed fitted values vs. untransformed residuals is shown in Appendix C. The detailed results of the regression model for the ARIMA(2,2,2) case are shown in Appendix B.

For the ARIMA(3,2,3) case, the results are again similar to the previous cases. The resulting regression equation is of the form

$$\begin{aligned} y^{-2.4} = {} & \beta_0 + \beta_1 x_1 x_2 + \beta_2 x_1 x_8 + \beta_3 x_2 x_3 + \beta_4 x_2 x_4 + \beta_5 x_2 x_6 + \beta_6 x_3 x_4 + \\ & \beta_7 x_3 x_6 + \beta_8 x_3 x_7 + \beta_9 x_3 x_8 + \beta_{10} x_5 x_6 + \beta_{11} x_6 x_7 + \beta_{12} x_6 x_8 \end{aligned} \tag{4.4}$$

previous two models, this R^2 is greatly affected by a few demand streams with large MOE values which are overestimated by the model. The decreased power of transformation used in the ARIMA(3,2,3) model results in a noticeably lessened distortion effect in the tails of the predicted values. The specific regression details for the ARIMA(3,2,3) model are shown in Appendix B.

All three of the regression models are highly predictive of their respective MOE measures. However, there is a general tendency to overpredict MOE measures when the actual MOE measure is above roughly 1.5. This should not be a significant problem in predicting ARIMA suitability, though, since any MOE value above 1.5 should be highly suggestive of ARIMA benefit to the time series.

There are, however, other shortcomings to regression analysis. A resulting regression model can be easily applied – but the development of the model is a detailed and complicated process. Due to the highly specialized nature of the model, an entirely new analysis would be required to consider suitability of other ARIMA models. A possible alternative to regression analysis is the use of statistical network software.

C. STATISTICAL NETWORK

This thesis uses the expert system statistical network software package ModelQuest™. ModelQuest™ develops polynomial-order networks of predictive variables to model a result variable (ModelQuest, 1995). In many aspects, ModelQuest™ performs a similar analysis to regression, however the model building process is entirely

performs a similar analysis to regression, however the model building process is entirely automated. The goal of using ModelQuest™ as an alternative to regression is to find an easier and faster method of generating new models, while maintaining predictive power.

ModelQuest™ determines both the network structure as well as the coefficients within the network. This is in contrast to regression, for which the network structure must be pre-defined. The additional range of freedom that ModelQuest™ has in selecting network structure is managed through a predicted squared error (PSE) equation (Barron, 1984). The PSE for a model is the sum of the fitted squared error (FSE) and a complexity penalty (KP). ModelQuest™ defines the optimal solution for a given situation as that one which minimizes the PSE value. This is shown schematically in Figure 3.

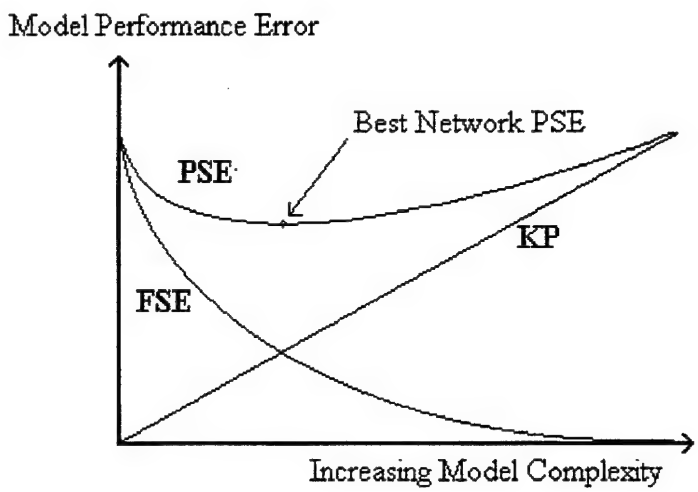


Figure 3: Relationship Between PSE, FSE, KP, and the Optimal Solution

FSE is comparable to the residual sum of squares for a particular regression model. ModelQuest™ defines KP in the following manner,

$$KP = (CPM)(\frac{2K}{N})(s_p^{\,2}) \tag{4.5}$$

where CPM is the complexity penalty multiplier
K is the number of coefficients within the network
N is the number of data elements being modeled
$s_p^{\,2}$ is an estimate of the model error variance

The value of the complexity multiplier (CPM) is user-definable. By adjusting the CPM, the complexity of a statistical network is controlled – an increased CPM results in a decreased complexity model.

In practice, managing CPM is essential to ensuring that the resulting model is neither over- or under-specific. This is done by comparing the performance of the model between two different datasets. The training dataset is used to develop the model, while the evaluation dataset is used to evaluate the developed model. If the performance of the model is significantly different between the training and evaluation datasets, then the model has overfit the training data. For this thesis, 75% of the entire database of features and MOE values were set aside for the training dataset and the remaining 25% formed the evaluation database. The selection of the particular elements for the training and evaluation sets was done randomly by ModelQuest™.

ModelQuest™ creates a layered network model. Polynomial relationships are created between variables within each layer of the network. These relationships may include the addition of constant terms, as well as polynomials of order up to three. A schematic network is shown in Figure 4.

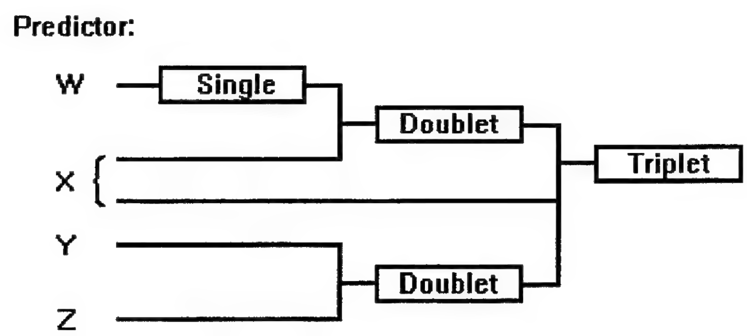


Figure 4: Example Three-Layer ModelQuest™ Network

At each "box" operator of the network in Figure 4, a polynomial of order three (or less) combines the left-side input variables of that box to form a result. Where an input variable is the output from a previous box, that previous polynomial is treated as if it were a single variable. In other words, the order of the polynomial equation which represents the entire network may be three times the size of n, where n is the total number of layers in the network. ModelQuest™ allows the definition of a strict upper bound for the number of layers within the network as well as the number of terms within the first layer of the network.

Other than CPM and the dataset itself, ModelQuest™ requires no other information or intervention to develop the network. The resulting networks are more complicated than regression results, as will be seen, but they have the benefit of completely automated formation. For the statistical network models in this thesis, several different values of CPM were tried for each network. The best network model is arrived at by comparing the performance within a model for the different datasets as well as comparing the performances between models with different CPM.

The structure of the statistical network model for the ARIMA(1,1,1) case is shown in Figure 5.

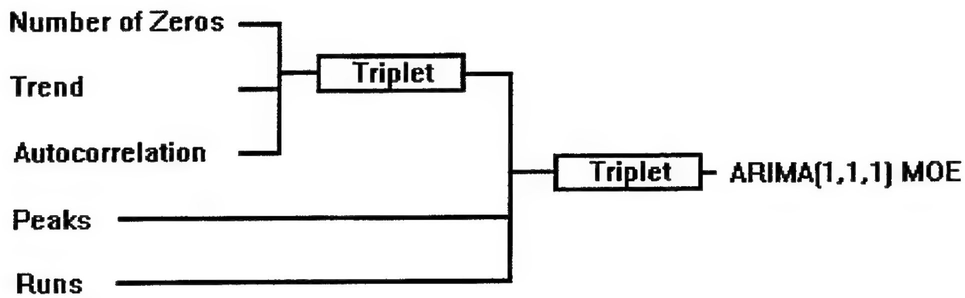


Figure 5: Statistical Network for ARIMA(1,1,1) Case

In comparison to the regression model, the statistical network uses only five of the eight features to predict MOE in this case. This is compensated for by the additional complexity of the sixth-order polynomial which results from two triplet nodes. The R^2 which results from this model is 0.8054 for the training dataset and 0.78032 for the evaluation dataset. Details of the statistical network as well as the diagnostics for this model are shown in Appendix D.

The statistical network for the ARIMA(2,2,2) case is more complex than the single order ARIMA model. This model uses six of the eight predictive features, adding the feature of seasonality to the five features used in the ARIMA(1,1,1) network. The structure of the model is shown in Figure 6.

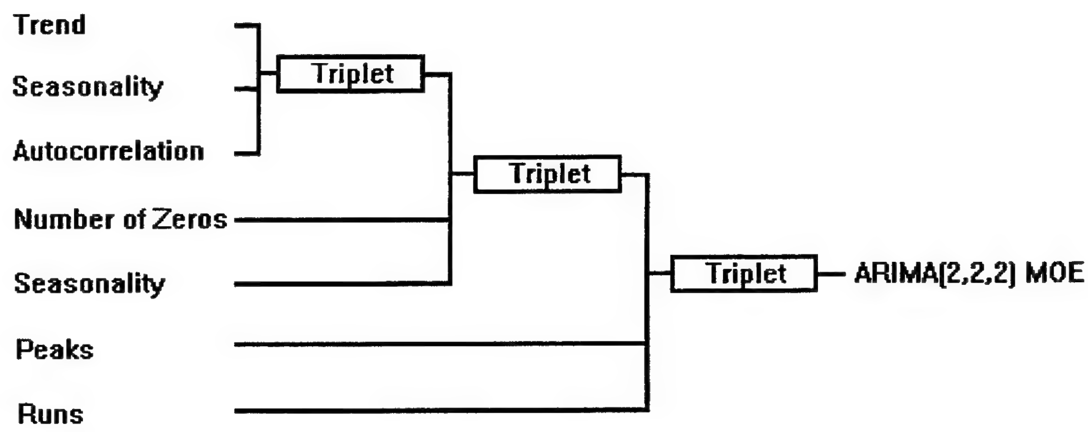


Figure 6: Statistical Network for ARIMA(2,2,2) Case

The structure of the ARIMA(2,2,2) network is very similar to that of the ARIMA(1,1,1) case. The primary difference between the structure of the two is the addition of a third triplet operation in the ARIMA(2,2,2) case. For the second order ARIMA case, the R^2 value is 0.5861 for the training set and 0.5869 for the evaluation set. The detailed form of the statistical network and the diagnostics of the model are shown in Appendix D.

Interestingly, the structure of the statistical network for the ARIMA(3,2,3) case is identical to the ARIMA(1,1,1) case. The only difference between the two networks are the coefficients of the predictors. For the third order ARIMA model, an R^2 of 0.5751 for the training set and 0.6131 for the evaluation set results. The detailed form of the network along with the diagnostics of the third order model are shown in Appendix D.

Applying the network structure of Figure 6 to the second order ARIMA case gives an R^2 of 0.5097. This provides strong evidence that there is a positive correlation between

MOE improvement among the ARIMA processes. Also, the structure shown in Figure 6 seems to be a good general structure for modeling Box-Jenkins suitability.

Both the ModelQuest™ and regression models exhibit a high predictive power of ARIMA MOE given the developed features. The choice as to which to use in actual application depends on the situation. ModelQuest™ requires only a minimum of training and a standard personal computer, however the results are not as detailed as those available in post-regression analysis. The regression process, in comparison, requires careful analysis at each step in the process. The results from regression are extremely detailed. However, both regression and ModelQuest™ provide excellent results when properly applied.

D. ARIMA FORECASTING

It is difficult to develop a meaningful measure of the improvement provided by ARIMA modeling over exponential smoothing in actual forecasting. The large variety of time series present in the NAVICP database complicate forecast comparisons. Some simple forecasts are made to quantify the ARIMA benefit. The first $n-1$ quarters of data are used to forecast the n^{th} quarter using both exponential smoothing and the ARIMA models. This is done for those time series from the analysis dataset which have an ARIMA MOE coefficient equal or greater than 1.25. The statistic that is measured to quantify improvement is:

$$\frac{\left|F_n^E - X_n\right|}{\left|F_n^A - X_n\right|} \tag{4.6}$$

where F_n^A is the ARIMA forecast for period *n*, F_n^E is the exponential smoothing forecast for period *n*, and X_n is the actual demand for period *n*

This ratio measures the improvement in forecasting from a Box-Jenkins model compared to exponential smoothing. The median improvement measure is: 1.10 for the ARIMA(1,1,1) model, 1.17 for the ARIMA(2,2,2) model, and 1.22 for the ARIMA(3,2,3) model. These results indicate that there is improvement to be gained from forecasting with Box-Jenkins models.

Forecasting from ARIMA models is noteworthy in that predicted demands of zero are possible. Exponential smoothing is incapable of producing a forecast of zero, however the ARIMA models produce these forecasts recurringly for the NAVICP data. These zero forecasts validate to be important due to the large number of repair parts characterized by frequent zero demand.

It is estimated that roughly 8% of NAVICP stock numbered parts would benefit from ARIMA modeling. Although this is a seemingly small fraction, it represents a more sizeable percentage of those repair parts which have recurring positive demand. These are the items which make up the vast majority of all demands, and will correspondingly provide a larger overall benefit with improved forecasts.

V. CONCLUSION

This thesis developed eight features to assist in classifying repair part demand streams. The features are notable in their ease of computation, robustness, and lack of dimension. When taken compositely, the features provide a means of predicting Box-Jenkins modeling suitability.

A. RESULTS

Both classical regression and expert system software were used to relate the time series features to the ARIMA suitability MOE. While arriving at different networks by different means, these two methods both result in a high predictability of Box-Jenkins suitability. By presenting two different approaches, there is versatility in implementation of the ARIMA suitability methodology. The repair parts which will likely benefit from Box-Jenkins modeling can be quickly identified through the models developed within the thesis.

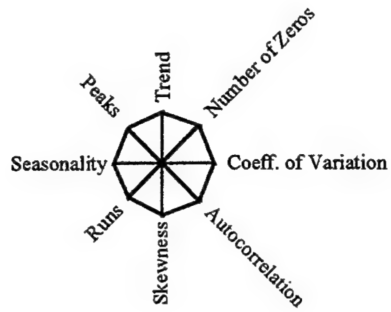
B. AREAS FOR FURTHER STUDY

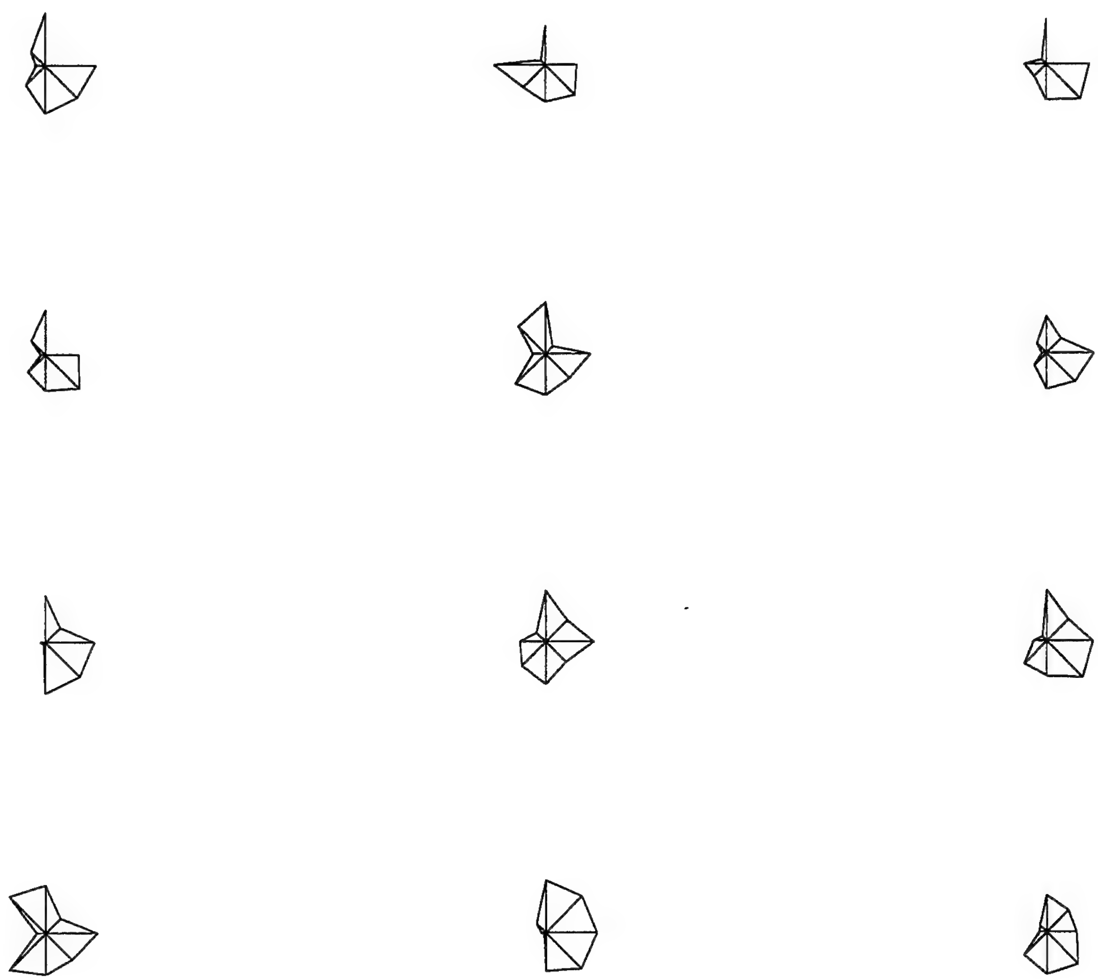
It would be beneficial to quantify the exact benefits of ARIMA modeling and forecasting for Navy repair part data. Such a study would possibly include a look into which types of repair parts receive benefit from Box-Jenkins and what the result of an improved forecasting technique is. Intuitively, an improvement in forecast accuracy leads to decreased risk of a stock out which will allow the safety stock level to be decreased.

Further research into demand stream characteristics may also provide interesting results. This thesis has developed several features, however additional features likely exist as well. A study of the processes within the high coefficient of variation subset would directly add to the work of this thesis. For the entire dataset as well, additional features would enhance both classification and ARIMA predictive power.

Sensitivity analysis of the ARIMA models may also be conducted as part of a comparison of the efficacy of different models. The versatility of Box-Jenkins through the different orders of autoregression, integration, and moving average as applied to the NAVICP data is unexplored. There are certainly tradeoffs between different models, and research into what they are and how significant a role they play would prove beneficial.

APPENDIX A. SAMPLE STAR PLOT DIAGRAMS





APPENDIX B. REGRESSION RESULTS AND PLOTS

ARIMA(1,1,1) case:

Formula: X9 ^ (-3.5) ~ X3 + X5 + X6 + X1*X3 + X2*X3 + X2*X5 + X2*X8 + X3*X8
X4*X7 + X5*X6

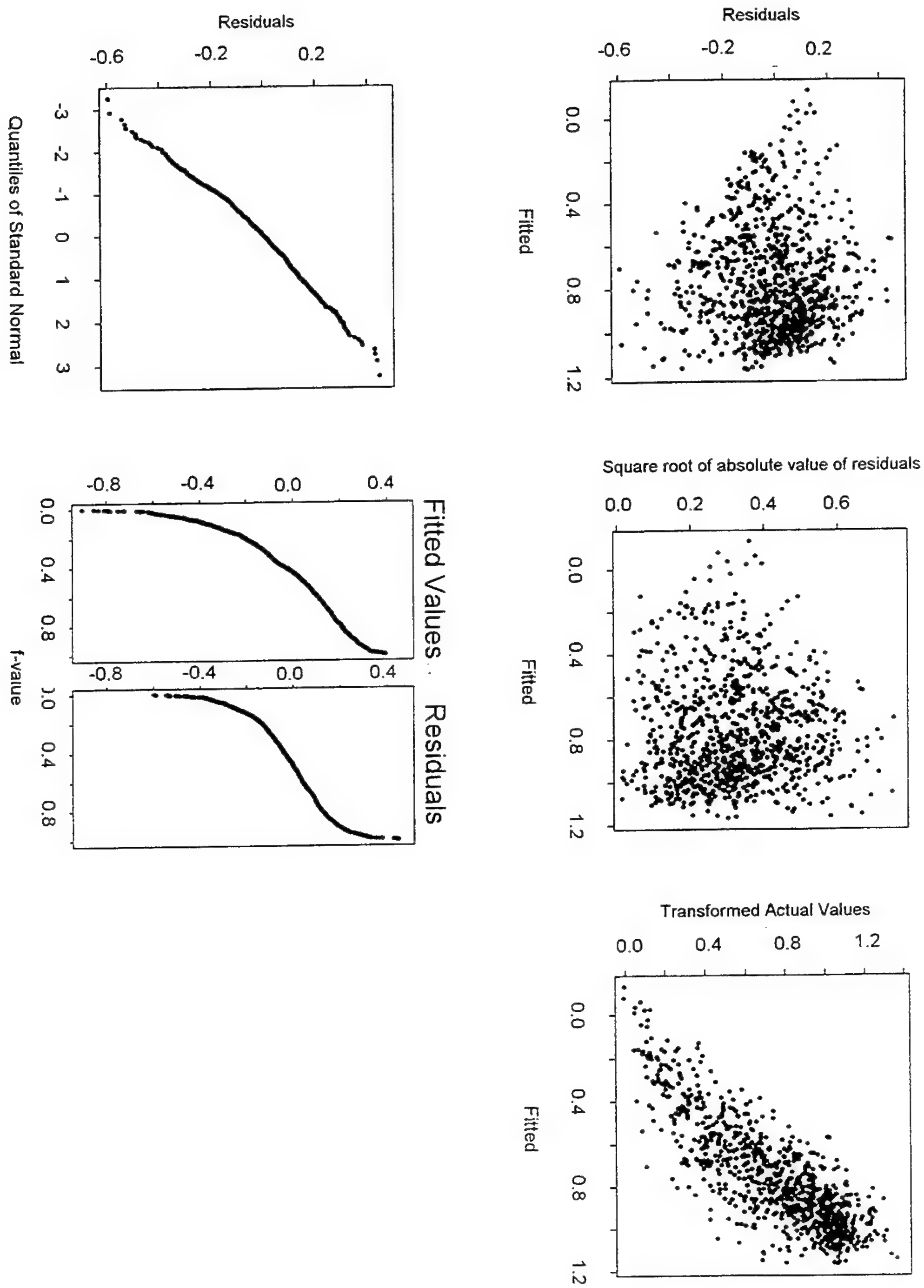
Residuals:

Min	1Q	Median	3Q	Max
-0.5841	-0.09582	0.01352	0.1115	0.4589

Coefficients:

	Value	Std. Error	t value	Pr(>|t|)
(Intercept)	0.6096	0.0383	15.9132	0.0000
X3	-0.3321	0.0747	-4.4476	0.0000
X5	1.7955	0.4373	4.1063	0.0000
X6	0.0866	0.0064	13.4368	0.0000
X1:X3	0.1534	0.0652	2.3525	0.0189
X2:X3	0.0154	0.0036	4.3194	0.0000
X2:X5	0.0581	0.0159	3.6501	0.0003
X2:X8	-0.0012	0.0005	-2.3047	0.0214
X3:X8	-0.1059	0.0124	-8.5294	0.0000
X4:X7	-0.0043	0.0013	-3.4195	0.0007
X5:X6	-0.3533	0.0765	-4.6191	0.0000

Residual standard error: 0.1689 on 867 degrees of freedom
Multiple R-Squared: 0.6848
F-statistic: 189.5 on 10 and 872 degrees of freedom, the p-value is 0


Residuals
Quantiles of Standard Normal
Residuals
Fitted
Fitted Values
Residuals
f-value
Square root of absolute value of residuals
Fitted
Transformed Actual Values
Fitted

ARIMA(2,2,2) case:

Formula: X9 ^ (-3.0) ~ X3 + X5 + X6 + X8 + X2*X3 + X3*X8 + X4*X7 + X5*X6 + X5*X8

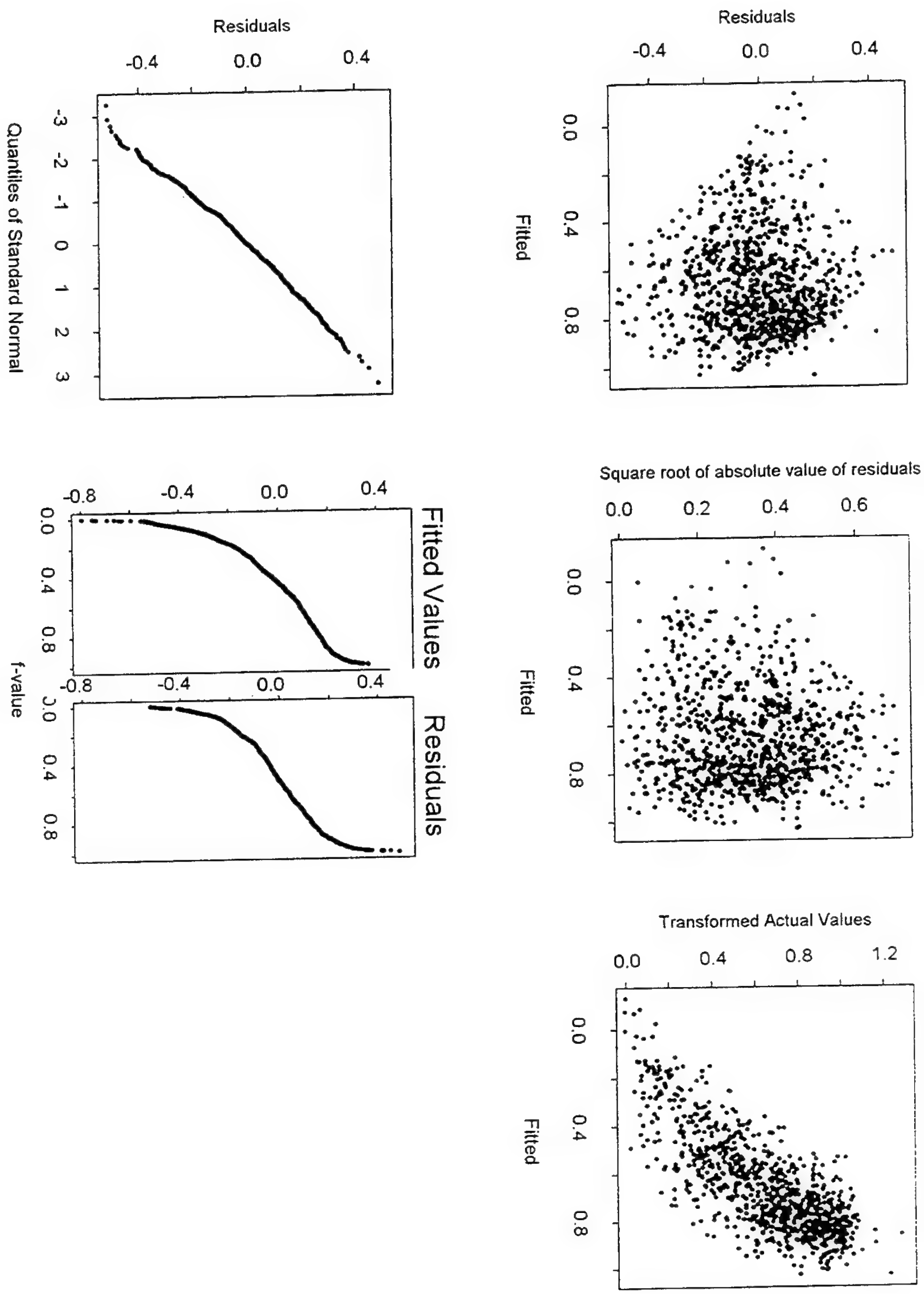
Residuals:

Min	1Q	Median	3Q	Max
-0.5093	-0.09787	0.003696	0.117	0.4963

Coefficients:

	Value	Std. Error	t value	Pr(>|t|)
(Intercept)	0.5414	0.0472	11.4773	0.0000
V3	-0.1713	0.0564	-3.0376	0.0025
V5	1.2312	0.5225	2.3565	0.0187
V6	0.0844	0.0069	12.1485	0.0000
V8	-0.0196	0.0122	-1.6056	0.1087
V2...V3	0.0120	0.0019	6.4248	0.0000
V3...V8	-0.0859	0.0185	-4.6574	0.0000
V4...V6	-0.0017	0.0004	-3.7216	0.0002
V5...V6	-0.3762	0.0789	-4.7656	0.0000
V5...V8	0.3168	0.0908	3.4904	0.0005

Residual standard error: 0.1676 on 864 degrees of freedom
Multiple R-Squared: 0.6071
F-statistic: 148.3 on 9 and 864 degrees of freedom, the p-value is 0


Residuals
Quantiles of Standard Normal
Fitted
Fitted Values
Residuals
f-value
Square root of absolute value of residuals
Transformed Actual Values

ARIMA(3,2,3) case:

Formula: X9 ^ (-2.4) ~ X1*X2 + X1*X8 + X2*X3 + X2*X4 + X2*X6 + X3*X4 + X3*X6 + X3*X7 + X3*X8 + X5*X6 + X6*X7 + X6*X8

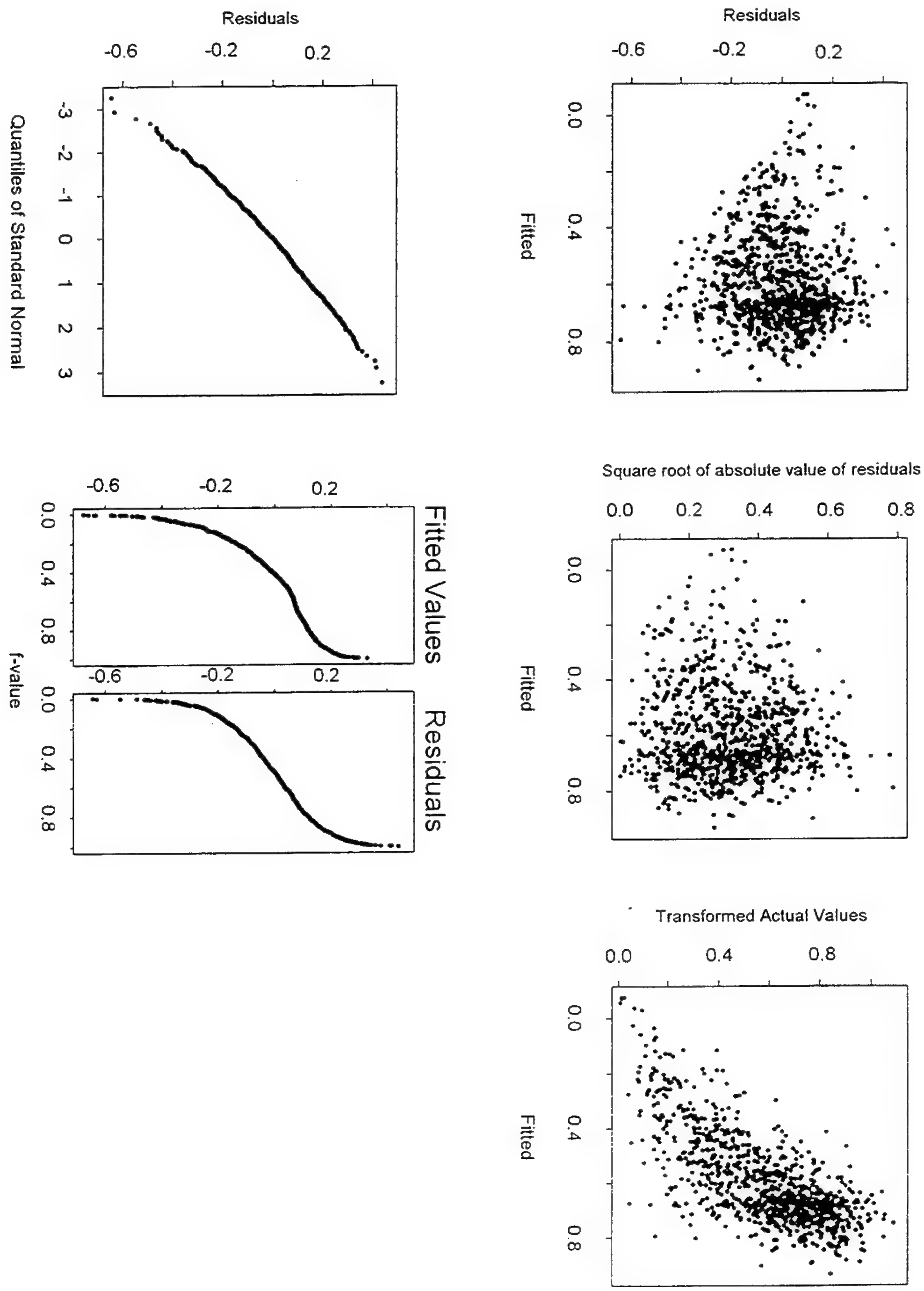
Residuals:

Min	1Q	Median	3Q	Max
-0.6269	-0.1018	0.0113	0.1024	0.4477

Coefficients:

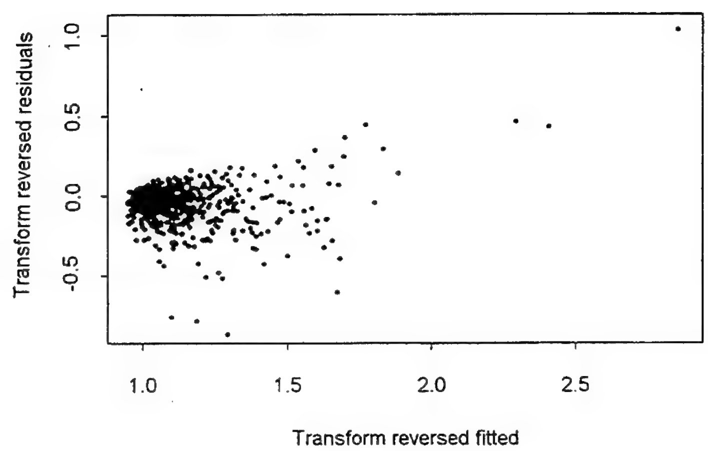
	Value	Std. Error	t value	Pr(>|t|)
(Intercept)	0.6457	0.0200	32.2421	0.0000
V1...V2	0.0052	0.0039	1.3483	0.1779
V1...V8	-0.0174	0.0145	-1.1954	0.2323
V2...V3	0.0073	0.0036	2.0326	0.0424
V2...V4	-0.0004	0.0004	-1.0806	0.2802
V2...V6	-0.0002	0.0007	-0.2698	0.7873
V3...V4	-0.0087	0.0046	-1.8947	0.0585
V3...V6	0.0159	0.0094	1.6863	0.0921
V3...V7	-0.0240	0.0183	-1.3160	0.1885
V3...V8	-0.1124	0.0133	-8.4247	0.0000
V5...V6	0.0212	0.0184	1.1502	0.2504
V6...V7	0.0042	0.0030	1.4131	0.1580
V6...V8	0.0085	0.0026	3.2751	0.0011

Residual standard error: 0.1592 on 866 degrees of freedom
Multiple R-Squared: 0.5337
F-statistic: 82.59 on 12 and 866 degrees of freedom, the p-value is 0

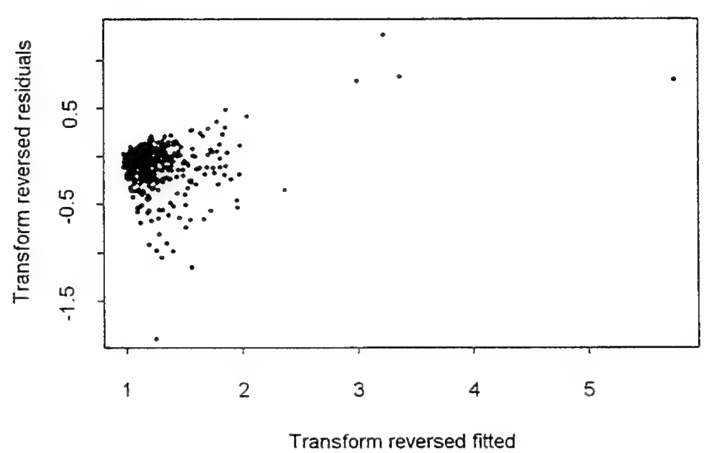

Residuals
Quantiles of Standard Normal
Fitted Values
Residuals
f-value
Residuals
Fitted
Square root of absolute value of residuals
Fitted
Transformed Actual Values
Fitted

APPENDIX C. UNTRANSFORMED REGRESSION RESIDUAL PLOTS

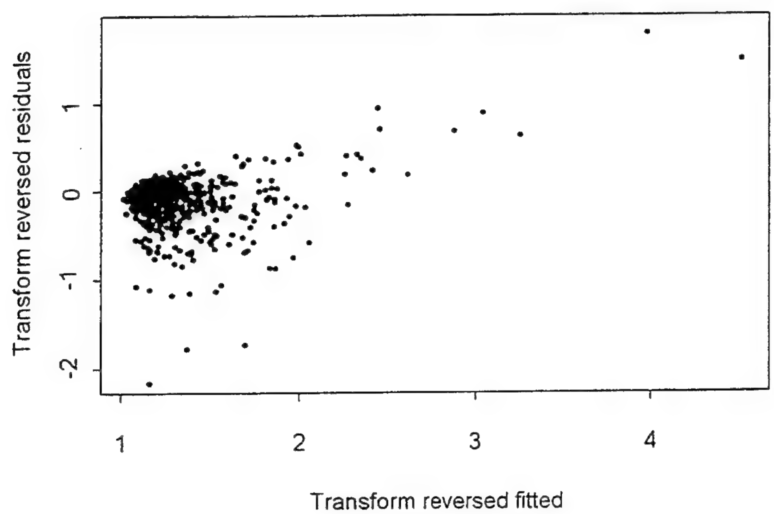
ARIMA(1,1,1) Regression



ARIMA(2,2,2) Regression



ARIMA(3,2,3) Regression



APPENDIX D. STATISTICAL NETWORK RESULTS

Summary Statistics for output 'ARIMA(1,1,1) MOE'

	evaluation	training
number of observations	221	662
average absolute error	0.070402	
absolute error standard deviation	0.067067	
average squared error	0.0094340	0.013497
normalized mean squared error	0.22933	
squared error standard deviation	0.019915	
maximum absolute error	0.41618	
database output minimum	0.91301	0.91883
database output maximum	2.2221	4.2954
database output mean	1.1318	1.1402
database output standard deviation	0.20327	0.26278
network output mean	1.1344	
network output standard deviation	0.19955	
R-squared	0.78032	
root of predicted squared error		0.15197
predicted squared error		0.023094

Summary Statistics for output 'ARIMA(2,2,2) MOE'

	evaluation	training
number of observations	218	656
average absolute error	0.13389	
absolute error standard deviation	0.15605	
average squared error	0.042165	0.025480
normalized mean squared error	0.34533	
squared error standard deviation	0.11735	
maximum absolute error	1.0232	
database output minimum	0.93120	0.91780
database output maximum	4.1915	4.7695
database output mean	1.2384	1.2273
database output standard deviation	0.34928	0.30813
network output mean	1.2127	
network output standard deviation	0.26087	
R-squared	0.65635	
root of predicted squared error		0.20524
predicted squared error		0.042124

Summary Statistics for output 'ARIMA(3,2,3) MOE'

	evaluation	training
number of observations	220	659
average absolute error	0.18444	
absolute error standard deviation	0.24832	
average squared error	0.095399	0.054653
normalized mean squared error	0.42826	
squared error standard deviation	0.39671	
maximum absolute error	2.1067	
database output minimum	0.96086	0.97399
database output maximum	5.2370	3.9581
database output mean	1.3583	1.3359
database output standard deviation	0.47302	0.35894
network output mean	1.3531	
network output standard deviation	0.27410	
R-squared	0.61307	
root of predicted squared error		0.26048
predicted squared error		0.067850

LIST OF REFERENCES

Barron, A.R., "Predicted Squared Error: A Criterion for Automatic Model Selection", *Self-Organizing Methods in Modeling: GMDH Type Algorithms*, edited by Farlow, S.J., Marcel-Dekker, Inc., 1984

Bissinger, B.H. and Boyarski, J.R., *Classification Strategies for Detection and Recognition of Patterns in Military Spare Parts Demand*, unpublished, 1992

Box, G.E. and Cox, D.R., "An analysis of transformations (with discussion)", *Journal of the Royal Statistical Society, Ser. B*, 1964

Box, G.E. and Jenkins, G.M., *Time Series Analysis: Forecasting and Control*, Holden-Day, Inc., 1970

Boyarski, J.R., *Notes on Repair Parts Classification*, unpublished, 1995

Chambers, J.M., Cleveland, W.S., Kleiner, B., and Tukey, P.A., *Graphical Methods for Data Analysis*, Wadsworth & Brooks/Cole Publishing Co., 1983

Gibbons, J.D. and Chakraborti, S., *Nonparametric Statistical Inference*, Marcel Dekker, Inc., 1992

Hoff, J.C., *A Practical Guide to Box-Jenkins Forecasting*, Wadsworth, Inc., 1983

McNeil, D.R., *Interactive Data Analysis*, John Wiley & Sons, Inc., 1977

Miller, R.G., *Beyond ANOVA, Basics of Applied Statistics*, John Wiley & Sons, Inc., 1986

ModelQuest, *ModelQuest User's Manual*, AbTech Corp., Inc., 1995

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center 2
   8725 John J. Kingman Rd., STE 0944
   Ft. Belvior, VA 22060-6218

2. Dudley Knox Library 2
   Naval Postgraduate School
   411 Dyer Rd.
   Monterey, CA 93943-5101

3. Defense Logistics Studies Information Exchange 1
   U.S. Army Logistics Management College
   Fort Lee, VA 22134-5130

4. Prof. R. Read, Code OR/RE 2
   Naval Postgraduate School
   Operations Analysis Department
   Monterey, CA 93943-5101

5. Prof. R. Whitaker, Code OR/WH 2
   Naval Postgraduate School
   Operations Analysis Department
   Monterey, CA 93943-5101

6. CDR R. Vassian, Code OR/BB 1
   Naval Postgraduate School
   Operations Analysis Department
   Monterey, CA 93943-5101

7. Mr. J. Engelman, Code M041X 2
   Navy Inventory Control Point
   5450 Carlisle Point
   Mechanicsburg, PA 17055

8. Mr. J. Boyarski, Code 0412 1
   Navy Inventory Control Point
   5450 Carlisle Point
   Mechanicsburg, PA 17055

9. LCDR(Sel) Mark Businger 7
   USS Valley Forge (CG-50)
   FPO AP 96682-1170